Response to Fee Exemption Request:

In support of our research project to develop a classifier for identifying hate speech, sexual harassment, and other nefarious acts within organizational communication systems, we plan to review the following documents from PACER. While we may not have the exact number of each document type, we can provide an estimate:

1. **Cases and Dockets:**
   - **Civil Rights Cases** (Suit Codes: 440, 441, 442, 443, 444)
     - Estimated Number: 500 cases
     - Courts: U.S. District Courts nationwide, with a focus on the Northern District of California, Southern District of New York, and District of Columbia.
   - **Employment Discrimination Cases** (Suit Codes: 442, 445)
     - Estimated Number: 400 cases
     - Courts: U.S. District Courts in California, New York, Texas, and Illinois.
   - **Harassment and Defamation Cases** (Suit Codes: 320, 370)
     - Estimated Number: 300 cases
     - Courts: U.S. District Courts in Florida, Pennsylvania, and Ohio.
   - **Hate Crimes and Related Incidents** (Suit Codes: 440, 443)
     - Estimated Number: 200 cases
     - Courts: U.S. District Courts in Michigan, Massachusetts, and Virginia.
2. **Dockets and Court Orders:**
   - **Dockets from Civil Rights and Employment Discrimination Cases**
     - Estimated Number: 1,500 dockets
     - Courts: Same as above
   - **Court Orders and Judgments in Relevant Cases**
     - Estimated Number: 800 documents
     - Courts: Same as above
3. **Specific Documents within Cases:**
   - **Complaints and Pleadings**
     - Estimated Number: 1,000 documents
     - Courts: U.S. District Courts in major jurisdictions as listed above.
   - **Motions and Briefs Related to Discrimination and Harassment**
     - Estimated Number: 700 documents
     - Courts: Same as above.
   - **Deposition Transcripts and Exhibits**
     - Estimated Number: 600 documents
     - Courts: Same as above.

We have carefully selected the U.S. District Courts as they stock the types of civil and harassment-related cases relevant to our research. If necessary, we will review and adjust our list of courts to ensure appropriate designation.

Research Topic and Definition of "Classifier":

Our research focuses on developing an advanced machine learning classifier designed to help organizations identify and address hate speech, sexual harassment, and other harmful acts within their controlled communication systems, such as email, Slack, etc. A "classifier" in this context refers to a software tool that uses machine learning algorithms to categorize and identify instances of specified behaviors based on the data it has been trained on.

Data Automation and Use of Scripts:

We plan to use data automation and scripts to efficiently process and analyze the large volume of documents from the PACER system. These scripts will not interact directly with the court's database system but will be used to organize, preprocess, and extract relevant information from the downloaded documents. This approach is crucial for handling the extensive data set required for training our classifier accurately and effectively.

We understand the importance of adhering to PACER's guidelines and ensuring the ethical use of court data. Therefore, we will ensure that our methods comply with all relevant regulations and standards.